第三章 矩阵分析和线性矩阵不等式

本章介绍本书中所用到的必要矩阵论方面数学知识，集中建立和证明一些常用的矩阵方程，不等式以及与线性矩阵不等式(LMI)有关的基本公式，这些公式与引理在后续各章中均有引用。

3.1 矩阵运算基础

3.1.1 矩阵运算

本小节给出矩阵运算中得一些基本定理和算法。

定理 3.1.1 考虑一般的$n+m$阶分块方阵

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{3.1.1}$$

式中A_{11}和A_{22}分别为n阶和m阶可逆矩阵，则有如下两个矩阵求逆公式，即

$$[A_{11}-A_{12}A_{22}^{-1}A_{21}]^{-1}=A_{11}^{-1}+A_{11}^{-1}A_{12}[A_{22}-A_{21}A_{11}^{-1}A_{12}]^{-1}A_{21}A_{11}^{-1} \tag{3.1.2}$$

和

$$A_{11}^{-1}A_{12}[A_{22}-A_{21}A_{11}^{-1}A_{12}]^{-1}=[A_{11}-A_{12}A_{22}^{-1}A_{21}]^{-1}A_{12}A_{22}^{-1} \tag{3.1.3}$$

证明：略。

定理 3.1.2 **（矩阵求逆引理）**

$$[A+BC]^{-1}=A^{-1}-A^{-1}B[I+CA^{-1}B]^{-1}CA^{-1} \tag{3.1.4}$$

证明：在(3.1.2)式中，令$A_{11}=A,\ A_{12}=B,\ A_{21}=C,\ A_{22}=I$，便得(3.1.4)式。

定理 3.1.3 如果A是按式(3.1.1)分块划分，则有

$$\det A=\det A_{22}\cdot\det[A_{11}-A_{12}A_{22}^{-1}A_{21}] \tag{3.1.5}$$

和

$$\det A=\det A_{11}\cdot\det[A_{22}-A_{21}A_{11}^{-1}A_{12}] \tag{3.1.6}$$

证明：因为

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & A_{22}^{-1} \end{bmatrix}=\begin{bmatrix} A_{11}-A_{12}A_{22}^{-1}A_{21} & A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \tag{3.1.7}$$

所以有

$$\det A\cdot\det A_{22}^{-1}=\det[A_{11}-A_{12}A_{22}^{-1}A_{21}]$$

注意到$\det A_{22}^{-1}=\dfrac{1}{\det A_{22}}$，便得(3.1.5)式。类似地可证(3.1.6)式。

定理 3.1.4 $\det(I+BC)=\det(I+CB)$ (3.1.8)

证明：考虑矩阵$A=\begin{bmatrix} I & B \\ C & I \end{bmatrix}$，将$A$分别代入(3.1.5)和(3.1.6)式便可证明其结果。

定理 3.1.5 $[I+BC]^{-1}B=B[I+CB]^{-1}$ (3.1.9)

证明：由(3.1.4)得

$$\begin{aligned}(I+BC)^{-1}B&=B-B(I+CB)^{-1}CB=B(I-(I+CB)^{-1}CB)\\&=B(I+CB)^{-1}(I+CB-CB)=B(I+CB)^{-1}\end{aligned}$$

矩阵的迹$tr[A]=\sum_{i=1}^{n}a_{ii}$也是一个非常重要得概念，下面是部分常用到得计算公式。

① $tr[A]=tr[A^T]$

② $tr[A_1+A_2]=tr[A_1]+tr[A_2]$

③ $tr[AB]+tr[BA]$

④ $tr[Axy^T]=tr[y^TAx]=y^TAx=x^TA^Ty$

⑤ 若A为同幂矩阵，则$tr[A]=rankA$

3.1.2 矢量与矩阵的微分运算

在鲁棒控制理论和系统建模中，矢量与矩阵的微分运算是非常重要的。本节我们不加证明地给出一些常用到得运算定理和公式。为了叙述方便，采用下列约定。

1） 标量对向量求导定义为行向量。若J为标量，$\theta\in R^n$，则

$$\frac{\partial J}{\partial\theta}=[\frac{\partial J}{\partial\theta_1},\frac{\partial J}{\partial\theta_2},\cdots,\frac{\partial J}{\partial\theta_n}] \tag{3.1.10}$$

2） 向量对向量求导定义为矩阵。若$x\in R^m$，$\theta\in R^n$，则

$$\frac{\partial x}{\partial\theta}=\begin{bmatrix}\frac{\partial x_1}{\partial\theta_1} & \cdots & \frac{\partial x_1}{\partial\theta_n}\\ \vdots & & \vdots\\ \frac{\partial x_m}{\partial\theta_1} & \cdots & \frac{\partial x_m}{\partial\theta_n}\end{bmatrix} \tag{3.1.11}$$

3） 矩阵对向量求导定义为增维矩阵。若$A\in R^{n\times n}$，$\theta\in R^n$，则

$$\frac{\partial A}{\partial\theta}=[\frac{\partial A}{\partial\theta_1},\frac{\partial A}{\partial\theta_2},\cdots,\frac{\partial A}{\partial\theta_n}] \tag{3.1.12}$$

4） 标量对矩阵求导仍为矩阵。设J为标量，M为矩阵，则$\frac{\partial J}{\partial M}$是以$\frac{\partial J}{\partial m_{ij}}$为第$ij$元素的矩阵，其中$m_{ij}$表示$M$矩阵的第$ij$元素。

在上述约定下，有如下一些结果：

1） $\frac{\partial}{\partial x}(a^Tx)=a^T$；

2） $\frac{\partial}{\partial x}(x^TAx)=2x^TA,\quad A$为对称矩阵；

3） $\frac{\partial}{\partial M}\log\det M=M^{-T}$；

4） $\frac{\partial M^{-1}}{\partial x}=-M^{-1}\frac{\partial M}{\partial x}M^{-1}$；

5） $\frac{\partial}{\partial x}tr[Ax^{-1}B]=-(x^{-1}BAx^{-1})^T$。

3.1.3 矩阵的最大秩分解与正交补的计算公式

如果矩阵 $A \in R^{m\times n}, A \neq 0$，令 $rank(A) = r(r>0)$， 那么总可以找到矩阵 $B \in R^{m\times r}$ 和 $C \in R^{r\times n}$，使得：

$$A = BC \tag{3.1.13}$$

成立，且

$$rank(A) = rank(B) = rank(C) \tag{3.1.14}$$

上式被称为最大秩分解。可以证明，矩阵的最大秩分解总是存在的。

令矩阵 $A \neq 0$，可以分解为：

$$A = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \begin{bmatrix} P_{r\times r} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \tag{3.1.15}$$

式中 U_1 为 A 阵不为零特征值对应的特征向量，U_2 为 A 阵零特征值对应的特征向量。那么 A 阵的正交补矩阵可由下式得出

$$A^{\perp} = TU_2 \tag{3.1.16}$$

其中 T 为任意的满秩矩阵。利用该方法可以得到矩阵

$$\begin{bmatrix} P \\ 0 \end{bmatrix}^{\perp} = [0 \quad Q] \tag{3.1.17}$$

其中 Q 是满足 $Q^T Q > 0$ 的任意矩阵。

3.1.4 矩阵的广义逆与线性方程组的解

对任意矩阵 A 都有其广义逆 A^- 存在，且满足：

$$AA^- A = A \tag{3.1.18}$$

上式一般都作为矩阵广义逆的定义公式。一般来说，对于任意矩阵的广义逆存在但并不唯一，且有：

$$rank(A^-) \geq rank(A) \tag{3.1.19}$$

成立。

根据前面的论述，任意矩阵 A 都有最大秩分解 $A = BC$，那么可以证明由下式定义的矩阵是 A 阵的广义逆矩阵，而且唯一。

$$A^+ = C^T(CC^T)^{-1}(B^T B)^{-1} B^T \tag{3.1.20}$$

由于该逆比较重要，所以单独命名为伪逆矩阵或称之为 Moore-Penrose 伪逆矩阵。这种伪逆矩阵在后面的滤波分析章节中有应用。

考虑线性矩阵方程组：

$$AX = Y$$

当 A 是非奇异矩阵时， $X = A^{-1}Y$ 是方程组的唯一解。

当 A 是奇异矩阵时，若 $rank(A,Y) = rank(A)$，称方程组是相容的，此时它有无数个解，解的通式为:

$$X = A^{-}Y + (I - A^{-}A)Z \qquad (3.1.21)$$

其中Z为任意矩阵向量。

当A是奇异矩阵时，且$rank(A,Y) \neq rank(A)$，称方程组是不相容的，此时线性矩阵方程组无解。但有使$\|AX - Y\|$最小的最小二乘解，而且该解始终存在，其通式为：

$$X = A^{+}Y + (I - A^{+}A)Z \qquad (3.1.22)$$

其中Z为任意矩阵向量。而且$X = A^{+}Y$是所有最小二乘解中使$X^{T}X$最小的唯一解。

3.1.5 奇异值及奇异值分解

本小节介绍矩阵的奇异值分解，它不仅是矩阵理论和矩阵计算的最基本和最重要的工具之一，而且在控制理论、系统辨识、信号处理等许多领域都被广泛地应用。

我们首先引入矩阵奇异值的定义，然后讨论它的许多性质。这里所有的引理和定理都不加证明，读者可以在许多参考书中找到。

一、奇异值分解

引理 3.1.1 设$A \in C^{m\times n}$，则

$$rank(A^{H}A) = rank(AA^{H}) = rank(A) \qquad (3.1.23)$$

引理 3.1.2 设$A \in C^{m\times n}$，则

(1) $A^{H}A$与AA^{H}的特征值均为非负实数；

(2) $A^{H}A$与AA^{H}的非零特征值相同，并且非零特征值的个数（重特征值按重数计算）等于$rank(A)$。

定义 3.1.1 设$A \in C^{m\times n}$，如果存在非负实数σ和非零向量$u \in C^{n}$，$v \in C^{m}$使得

$$Au = \sigma v, \qquad A^{H}v = \sigma u \qquad (3.1.24)$$

则称σ为A的奇异值，u和v分别称为A对应与奇异值σ的右奇异向量和左奇异向量。

由式(3.1.24)可得

$$A^{H}Au = \sigma A^{H}v = \sigma^{2}u \qquad (3.1.25)$$

$$AA^{H}v = \sigma Au = \sigma^{2}v \qquad (3.1.26)$$

因此σ^{2}是$A^{H}A$的特征值，也是AA^{H}的特征值，而u和v分别是$A^{H}A$和AA^{H}对应于特征值σ^{2}的特征向量。

定理 3.1.6 设$A \in C^{m\times n}$，且$rank(A) = r$，则存在m阶酉矩阵V和n阶酉矩阵U使得

$$V^{H}AU = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \qquad (3.1.27)$$

其中$\Sigma = diag(\sigma_{1},\cdots,\sigma_{r})$，且$\sigma_{1} \geq \cdots \geq \sigma_{r} > 0$

(3.1.27)称为矩阵A的奇异值分解。从定理 3.1.6 可见，U的列向量是$A^{H}A$的标准正交特征向量，U的前r列向量是$A^{H}A$对应于r个非零特征值$\sigma_{1}^{2},\cdots,\sigma_{r}^{2}$的标准正交特征向量；而$V$的列向量是$AA^{H}$的标准正交特征向量，从$V_{1} = AU_{1}\Sigma^{-1}$的选取，有

$$AA^{H}V_{1} = AA^{H}AU_{1}\Sigma^{-1} = AU_{1}\Sigma^{2}\Sigma^{-1} = AU_{1}\Sigma = V_{1}\Sigma^{2} \qquad (3.1.28)$$

这表明V的前r列向量恰是AA^H对应于特征值$\sigma_1^2,\cdots,\sigma_r^2$的标准正交特征向量。

二、矩阵奇异值的性质

设$A\in C^{m\times n}$，并且$\bar{\sigma}$表示最大奇异值，$\underline{\sigma}$表示最小奇异值，则有

（1）$A^H A\le \bar{\sigma}^2(A)*I_n,\quad AA^H\le \bar{\sigma}^2(A)*I_m$；

（2）$\underline{\sigma}(A)\le|\lambda_i(A)|<\bar{\sigma}(A)$；

（3）$\bar{\sigma}(A)=\max\limits_{x\neq 0}\dfrac{\|Ax\|}{\|x\|}=\max\limits_{\substack{x\neq 0\\ \|x\|=1}}\|Ax\|,\quad \underline{\sigma}(A)=\min\limits_{x\neq 0}\dfrac{\|Ax\|}{\|x\|}=\min\limits_{\substack{x\neq 0\\ \|x\|=1}}\|Ax\|$；

（4）$\bar{\sigma}(A)\le 1$当且仅当$I-A^H A\ge 0$或$I-AA^H\ge 0$；

（5）A满秩当且仅当$\underline{\sigma}(A)>0$；

（6）如果A可逆，则$\bar{\sigma}(A)=\underline{\sigma}(A^{-1})$；

（7）对于标量$\beta\in C$，有$\sigma_i(\beta A)=|\beta|\sigma_i(A)$；

（8）设$A_1,A_2\in C^{m\times n}$，则 $\bar{\sigma}(A_1+A_2)\le\bar{\sigma}(A_1)+\bar{\sigma}(A_2)$；

（9）设$A\in C^{m\times q}$，$B\in C^{q\times n}$则有 $\bar{\sigma}(AB)\le\bar{\sigma}(A)\bar{\sigma}(B)$和$\underline{\sigma}(AB)\ge\underline{\sigma}(A)\underline{\sigma}(B)$；

（10）设A可分块为$A=\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}$则有 $\bar{\sigma}(A)\ge\bar{\sigma}(A_{ij}),\quad i,j=1,2$。

三、奇异值与其它矩阵量的关系

(1) 范数

由矩阵的奇异值分解定理，并注意到矩阵A的L_2范数$\|A\|_2$是酉不变的，故

$$\|A\|_2=\left[\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right]^{1/2}=\|U^H AV\|=\|\Sigma\|_2=\sqrt{\sigma_1^2+\cdots+\sigma_r^2}\qquad(3.1.29)$$

即是说，任何一个矩阵的L_2范数等于该矩阵所有非零奇异值平方和的正平方根。

考虑矩阵A的秩k近似，并将其记作A_k，其中$k<r=rank(A)$。定义A_r如下：

$$A_r=\sum_{i=1}^{k}u_i\times\sigma_i\times v_i^H,\qquad k<r$$

则A与秩为k的任一矩阵B之差的L_1和L_2范数分别为

$$\min_{rank(B)=k}\|A-B\|_1=\|A-A_k\|_1=\sigma_{k+1}\qquad(3.1.30)$$

和

$$\min_{rank(B)=k}\|A-B\|_2^2=\|A-A_k\|_2^2=\sigma_{k+1}^2+\cdots+\sigma_r^2 \tag{3.1.31}$$

这一重要结果是许多概念和应用（例如总体最小二乘、数据压缩、图像增强、动态系统实现理论。以及 $Ax=b$ 的求解需要用一个低秩矩阵近似 A 的各类问题）的基础。

(2) 行列式

设 A 是 $n\times n$ 方矩阵。由于酉矩阵的行列式之绝对值等于 1，所以由定理 6.2.1 有：

$$|\det A|=\left|\det\sum\right|=\sigma_1\sigma_2\cdots\sigma_n \tag{3.1.32}$$

若所有 σ_i 都不等于零，则 $|\det A|\neq 0$，这意味着 A 是非奇异的；若至少有一个 $\sigma_i(i>r)$ 等于零，则 $\det A=0$，即 A 是奇异的。这就是之所以把全部 σ_i 值统称为奇异值的原因。综合式(3.1.31)和式(3.1.32)，对于一个 $n\times n$ 矩阵 A，下列不等式成立：

$$\begin{aligned}
&n\sigma_1\ge\|A\|_2\ge\sigma_1\\
&\sigma_1^n\ge\sigma_1^{n-1}\sigma_n\ge|\det A|\ge\sigma_n^n\\
&|\det A|^{1/n}\ge\sigma_n\ge|\det A|^{1/n}\\
&|\det A|^{1/n}\ge\sigma_n\ge|\det A|/\|A\|_2^{n-1}\\
&\|A\|_2^n/|\det A|\ge\sigma_1/\sigma_n\ge\max\left\{1,\frac{1}{n}\|A\|_2/|\det A|^{1/n}\right\}
\end{aligned} \tag{3.1.33}$$

这些不等式虽然是粗略的评价，但有时是有用的。

(3) 条件数

对于一个 $n\times n$ 矩阵 A，其条件数也可以利用奇异值定义为

$$cond(A)=\sigma_1/\sigma_h,\qquad h=\min(m,n) \tag{3.1.34}$$

由定义式(3.1.34)可以看出，条件数是一个大于或等于 1 的正数。条件数小的矩阵叫做“良性”矩阵，反之叫做“病态”矩阵。显然，一个奇异矩阵的条件数为无穷大，而条件数虽然不是无穷大，但却很大时，就称是接近奇异的；这意味着，在这种情况下的行向量或列向量的独立性很弱。另由 $cond(A)=\|A\|\|A^{-1}\|$ 易知，正交或酉矩阵 V 是“理想条件”的。式(3.1.34)也可用作 $cond(A)$ 的评价式。

考虑超定方程 $Ax=b$。此时，由于 A^HA 的奇异值分解为

$$A^HA=V\Sigma^2V^H \tag{3.1.35}$$

即矩阵 A^HA 的最大和最小奇异值分别是矩阵 A 的最大和最小奇异值的二倍，故

$$cond\left(A^HA\right)=\frac{\sigma_1^2}{\sigma_n^2}=[cond(A)]^2 \tag{3.1.36}$$

换言之，矩阵 A^HA 的条件数是矩阵 A 的条件数的平方倍。

(4) 特征值

设 $n\times n$ 方矩阵 A 的特征值为 $\lambda_1,\cdots,\lambda_n(|\lambda_1|>\cdots>|\lambda_n|)$，奇异值为

$\sigma_1,\cdots$，$\sigma_n\left(\sigma_1\geq\cdots\sigma_n\geq0\right)$，则

$$\sigma_1\geq\left|\lambda_i\right|\geq\sigma_n,\quad i=1,\cdots,n$$

$$cond\left(A\right)\geq\left|\lambda_1\right|/\left|\lambda_n\right|$$

特别值得指出的是，奇异值分解提供了在实际控制信号处理中发生的一些重要问题的定量答案。

问题 1：一矩阵与低秩矩阵是如何接近的？

答案：如果 $A\in C^{m\times n}$ 具有奇异值

$$\sigma_1\geq\cdots\geq\sigma_r\geq\sigma_{r+1}=\cdots=\sigma_n=0$$

则 $rank\left(A\right)=r$，且

$$\min_{rank\left(B\right)\leq r}\left\|A-B\right\|_2^2=\sigma_{r+1}^2$$

特别地，若 $A\in C^{n\times n}$ 非奇异，则 σ_n 是到奇异矩阵集合的距离。因此，$1/cond\left(A\right)=\sigma_n/\sigma_1=\sigma_n/\left\|A\right\|_2$ 是 A 接近奇异的相对测度。

问题 2：一矩阵的值域与零空间是什么？

答案：如果 $U=\left[u_1\cdots u_m\right]$ 和 $V=\left[v_1\cdots v_n\right]$ 分别是左和右奇异向量矩阵的列分块，并且 $rank\left(A\right)=r$，则 $null\left(A\right)=span\left\{v_{r+1},\cdots,v_n\right\}$ 和 $range\left(A\right)=span\left\{u_1,\cdots,u_r\right\}$。

问题 3：怎样度量两个 C^n 的 k 维子空间的接近？

答案：假定 $n\times k$ 矩阵 $Y=\left[y_1\cdots y_k\right]$ 和 $Z=\left[z_1\cdots z_k\right]$ 具有正交列，并且 $S_1=range\left(Y\right)$，$S_2=range\left(Z\right)$。若 $\sigma_1\geq\cdots\geq\sigma_k$ 是 Y^HZ 的奇异值，则

$$dist\left(S_1,S_2\right)=\min_{y\in S_1,z\in S_2,\left\|z\right\|_2=1}\left\|y-z\right\|_2=\sqrt{1-\sigma_1^2}$$

问题 4：具有共同零向量的两个矩阵是否接近？

答案：如果 A 和 B 均为 $m\times n$ 矩阵，并且 $\sigma_1\geq\cdots\geq\sigma_n$ 是 $C=\begin{bmatrix}A\\B\end{bmatrix}$ 的奇异值，则存在矩阵 E_A 和 E_B，它们具有以下性质：$\left\|E_A\right\|_2\leq\sigma_n$，$\left\|E_B\right\|\leq\sigma_n$,并且

$$null\left(A+E_A\right)\cap null\left(B+E_B\right)\neq\left\{0\right\}$$

3.2 线性矩阵不等式

20 世纪 70 年代初，Willems 首先将线性二次型最优控制问题转化为线性矩阵不等式求解，并且建立了线性矩阵不等式与矩阵 RICCATI 方程之间的关系。20 世纪 80 年代，H_∞ 控制问题的提出和研究，更加促进了线性矩阵不等式的研究和发展：第一是因为控制中的许多问题由于复杂性的增加，而

不可能直接给出问题求解的解析表达式，但是却可以将问题转化为 *LMI* 求解，从而可以利用现有的各种优化方法，特别是近年来发展起来的内点算法，求得问题的唯一最优解；第二是由于在工程实际问题中，对控制系统提出的性能指标要求可能是多方面的，利用传统的矩阵等式方程，求解系统控制问题只能得到满足某一方面性能指标的最优唯一解。但是，利用 *LMI* 由于可以将对控制系统提出的多目标性能指标转化为对 *LMI* 解集的多个凸约束，因此就可以很容易利用凸优化技术得到满足不同性能指标最优的各个解，即得到一个解集。所以，*LMI* 的求解在控制系统的分析、设计和系统辨识中扮演着一个重要的角色。

3.2.1 *LMI* 问题介绍

线性矩阵不等式具有如下的形式

$$F(x)=F_0+\sum_{i=1}^{m}x_iF_i>0 \tag{3.2.1a}$$

其中，$F_i=F_i^T\in R^{n\times n}$ 为给定的对称矩阵，$x\in R^m$ 为变量。

在(3.2.1)式中，$F(x)$ 为正定矩阵，即对于任何非零向量 $u\in R^n$ 有，$u^TF(x)u>0$ 成立。具有(3.2.1)式形式的线性矩阵不等式称为严格 LMI；当

$$F(x)\geq 0 \tag{3.2.1b}$$

时，称为非严格 LMI。

尽管式(3.2.1)具有一般性，但在控制理论中使用到的线性矩阵不等式却很少以这样的形式出现，而更多的是以矩阵形式出现，例如 Lyapunov 矩阵不等式

$$A^TX+XA<0 \tag{3.2.2}$$

如果上式中的系数矩阵 $A=\begin{bmatrix}-1 & 2\\ 0 & -2\end{bmatrix}$，变量 $X=\begin{bmatrix}x_1 & x_2\\ x_2 & x_3\end{bmatrix}$，则式(3.2.2)可以写成式(3.2.1)的形式

$$x_1\begin{bmatrix}-2 & 2\\ 2 & 0\end{bmatrix}+x_2\begin{bmatrix}0 & -3\\ -3 & 4\end{bmatrix}+x_3\begin{bmatrix}0 & 0\\ 0 & -4\end{bmatrix}<0$$

在控制理论的许多问题中，可以将式(3.2.1)中的变量 x_i 写成 X，而 X 可以是向量，也可以是矩阵，只要矩阵不等式线性依赖于变量 X，则矩阵不等式就称为 LMI。因此，无论在控制理论研究还是矩阵不等式的求解，我们都采用式(3.2.2)的矩阵形式。

利用 Schur 补可以将非线性不等式转换为 *LMI* 形式，其方法是，对于具有如下形式的 *LMI*

$$\begin{bmatrix}Q(x) & S(x)\\ S^T(x) & R(x)\end{bmatrix}>0 \tag{3.2.3}$$

等价于

$$R(x)>0,\quad Q(x)-S(x)R^{-1}(x)S^T(x)>0 \tag{3.2.4}$$

或者

$$Q(x)>0,\quad R(x)-S^T(x)Q^{-1}(x)S(x)>0 \tag{3.2.5}$$

成立。其中，$Q(x)=Q^T(x)$，$R(x)=R^T(x)$，以及 $S(x)$ 都是变量 x 的函数。相反，非线性不等式(3.2.4)和(3.2.5)式可以表示成(3.2.3)式形式的 LMI。

对于二次矩阵不等式可以向 LMI 转化：考虑二次矩阵不等式

$$A^T P + PA + PBR^{-1}B^T P + Q < 0$$

其中$A, B, Q = Q^T, R = R^T > 0$是给定的具有适当维数的矩阵，$P = P^T$是变量，该二次矩阵不等式可以表述为如下形式的$LMI$：

$$\begin{bmatrix} -A^T P - PA - Q & PB \\ B^T P & R \end{bmatrix} > 0 \tag{3.2.6}$$

定义 3.2.1（LMI 可行解性问题—feasible problem） 对于线性矩阵不等式(3.2.1)，如果存在x_i使得其成立，则称此LMI可解，简称为 FP 问题。

3.2.2 可化为LMI的问题

有关LMI问题的标准形式有以下几种：

(1) 凸规划可行性问题（$LMIP$）

给定LMI $F(x) > 0$，$LMIP$问题是求解x^{feas}使得

$$F(x^{feas}) > 0 \tag{3.2.7}$$

成立，或者是确定该LMI是否存在可行性解；

对于 Lyapunov 稳定性问题，LMI可行解可以描述为：给定$A_i \in R^{n\times n}$，$i = 1,2,\cdots,l$，寻找正定矩阵$P > 0$，满足LMI

$$P > 0, A_i^T P + PA_i < 0, i = 1,2,\cdots,l \tag{3.2.8}$$

或者不存在这样的矩阵。

(2) 特征值问题（EVP）

EVP 问题是在满足一定的LMI约束条件下，求解矩阵的最大最小特征值，即

$$\begin{aligned} &\min imize \quad \lambda \\ &subject \quad to \quad \lambda I - A(x) > 0, \quad B(x) > 0 \end{aligned} \tag{3.2.9}$$

其中，$A(x), B(x)$为对称矩阵，且是最优变量的x函数。此外，EVP 还可以作为求线性函数最小值问题的等价形式，即

$$\begin{aligned} &\min imize \quad c^T x \\ &subject \quad to \quad F(x) > 0 \end{aligned} \tag{3.2.10}$$

其中，$F(x)$为x的函数。

(3) 广义特征值问题（GEVP）

GEVP 问题是在满足一定的LMI约束条件下，求解矩阵的最大最小广义特征值问题，即

$$\begin{aligned} &\min imize \quad \lambda \\ &subject \quad to \quad \lambda B(x) - A(x) > 0, \quad B(x) > 0, \quad C(x) > 0 \end{aligned} \tag{3.2.11}$$

其中，$A(x), B(x), C(x)$为对称矩阵，且是最优变量的x函数。此外，GEVP 还有另外一种形式，

$$\begin{aligned} &\min imize \quad \lambda \\ &subject \quad to \quad A(x, \lambda) > 0 \end{aligned} \tag{3.2.12}$$

其中，$A(x,\lambda)$ 为当 x 或者 λ 为固定值时，分别为 x 或者 λ 的函数，并且满足单调性条件

$$\lambda > \mu \Rightarrow A(x,\lambda) > A(x,\mu) \tag{3.2.13}$$

(4) 凸优化问题（CP）

对称 $A(x), B(x)$ 是关于 x 的仿射函数，凸优化问题可以描述为

$$\begin{aligned} &\min imize \quad \log\det A(x)^{-1} \\ &subject \quad to \quad A(x) > 0, \quad B(x) > 0 \end{aligned} \tag{3.2.14}$$

当 $A > 0$ 时，$\log\det A^{-1}$ 是 A 的凸函数。凸优化问题可以转化为 EVP 问题，因为 $\det A(x) > \lambda$ 可以表达为关于 x 和 λ 的 LMI 。

3.2.3 求解线性矩阵不等式可行解的内点算法

由于求解 LMI 问题具有成熟的工具软件包 *MATLAB LMItoolbox*，该工具包在解决 LMI 问题时使用的是内点算法。内点算法自 1988 年开始就用于求解标准的 LMI 问题，为对内点算法加以介绍。

考虑 LMI ：$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i > 0$，其中 $F_i = F_i^T \in R^{n\times n}$ 为给定的对称矩阵，定义关于 $F(x)$ 的函数

$$\phi(x) = \begin{cases} \log\det F(x)^{-1}, F(x) > 0 \\ \infty, otherwise \end{cases} \tag{3.2.15}$$

是有限的，当且仅当 $F(x) > 0$。可行解集 $\{x \mid F(x) > 0\}$ 非空有界意味着矩阵 $F_1, F_2, \cdots, F_m$ 线性无关。

对标准的 EVP 问题

$$\begin{aligned} &\min imize \quad c^T x \\ &subject \quad to \quad F(x) > 0 \end{aligned}$$

假定其最优解为 λ^{opt}，对于所有的 $\lambda > \lambda^{opt}$，LMI

$$F(x) > 0, c^T x < \lambda$$

是可解的。

具体地求解步骤为：

(1) 给定初值 $\lambda^{(0)}, x^{(0)}$，使得 $F(x^{(0)}) > 0$ 以及 $c^T x^{(0)} < \lambda^{(0)}$；

(2) $\lambda^{(k+1)} = c^T x^{(k)} + \lambda^{(k)}$；

(3) $x^{(k+1)} = x^*(\lambda^{(k)})$

内点算法中所用的搜索方法为共轭梯度法，其具有两个优点，其一是与 QR 或者 Choslesky 方法相比，计算量小；其二是在达到收敛前有可能停止共轭梯度搜索，且仍然得到适合与内点方法的 Newton 直接方法的近似解。

3.3 LMI 中几个基本的矩阵变换公式

3.3.1 Schur 公式

1)给定矩阵$A \in R^{r\times r}, B \in R^{r\times(n-r)}, C \in R^{(n-r)\times r}, D \in R^{(n-r)\times(n-r)}$，并且方阵$A$可逆，则

$$(1)\quad \begin{bmatrix} I_{(r)} & 0 \\ -CA^{-1} & I_{(n-r)} \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & D-CA^{-1}B \end{bmatrix} \tag{3.3.1}$$

$$(2)\quad \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I_{(r)} & -A^{-1}B \\ 0 & I_{(n-r)} \end{bmatrix} = \begin{bmatrix} A & 0 \\ C & D-CA^{-1}B \end{bmatrix} \tag{3.3.2}$$

$$(3)\quad \begin{bmatrix} I_{(r)} & 0 \\ -CA^{-1} & I_{(n-r)} \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I_{(r)} & -A^{-1}B \\ 0 & I_{(n-r)} \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & D-CA^{-1}B \end{bmatrix} \tag{3.3.3}$$

2）给定矩阵$A \in R^{r\times r}, B \in R^{r\times(n-r)}, C \in R^{(n-r)\times r}, D \in R^{(n-r)\times(n-r)}$，并且方阵$D$可逆，则

$$(1)\quad \begin{bmatrix} I_{(r)} & -BD^{-1} \\ 0 & I_{(n-r)} \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A-BD^{-1}C & 0 \\ C & D \end{bmatrix} \tag{3.3.4}$$

$$(2)\quad \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I_{(r)} & 0 \\ -D^{-1}C & I_{(n-r)} \end{bmatrix} = \begin{bmatrix} A-BD^{-1}C & B \\ 0 & D \end{bmatrix} \tag{3.3.5}$$

$$(3)\quad \begin{bmatrix} I_{(r)} & -BD^{-1} \\ 0 & I_{(n-r)} \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I_{(r)} & 0 \\ -D^{-1}C & I_{(n-r)} \end{bmatrix} = \begin{bmatrix} A-BD^{-1}C & 0 \\ 0 & D \end{bmatrix} \tag{3.3.6}$$

3.3.2 Schur 补引理

定理 3.3.1 给定矩阵$Q \in R^{r\times r}, S \in R^{r\times(n-r)}, R \in R^{(n-r)\times(n-r)}$，

$$\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} < 0 \tag{3.3.7}$$

当且仅当

$$(\mathrm{i})\quad R<0, \qquad Q-SR^{-1}S^T<0 \tag{3.3.8}$$

或者

$$(\mathrm{ii})\quad Q<0, \qquad R-S^TQ^{-1}S<0 \tag{3.3.9}$$

成立。

证明：由于合同变换不改变矩阵的正定性。

先证明$\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} < 0$等价于$R<0$， $Q-SR^{-1}S^T<0$

当$Q<0$时，利用 Schur 公式的(3.3.3）式，有

$$\begin{bmatrix} I_{(r)} & -SR^{-1} \\ 0 & I_{(n-r)} \end{bmatrix} \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} I_{(r)} & 0 \\ -R^{-1}S^T & I_{(n-r)} \end{bmatrix} = \begin{bmatrix} Q-SR^{-1}S^T & 0 \\ 0 & R \end{bmatrix} \tag{3.3.10}$$

因此，有结论(3.3.7)成立，等价于(3.3.8)成立。

同理，可证(3.3.9)成立。

[证毕]

注：式中的$<$改为$\le$时，其结论依然成立，但此时R^{-1}应相应为R^{-}，或者Q^{-1}应相应为Q^{-}。在上述的公式中，$Q-SR^{-1}S^T$称为Q的Schur补。

3.3.3 消元公式

定理3.3.2 给定G,U,V，$G\in R^{n\times n}$，U,V具有n行，存在X以致于

$$G+UXV^T+(UXV^T)^T<0 \tag{3.3.11}$$

当且仅当

$$U^{\perp T}GU^{\perp}<0 \tag{3.3.12}$$

$$V^{\perp T}GV^{\perp}<0 \tag{3.3.13}$$

其中$U^{\perp},V^{\perp}$为U和V的正交补，有$U^{\perp}U=0,V^{\perp}V=0$。

证明：(1)必要性：

$U^{\perp T}U=0$ 意味着 $U^{\perp T}GU^{\perp}>0$。

因为对(3.3.11)式左乘$U^{\perp T}$和右乘$U^{\perp}$立即得到：

$$U^{\perp T}GU^{\perp}+U^{\perp T}UXV^TU^{\perp}+U^{\perp T}(UXV^T)^TU^{\perp}<0$$

即

$$U^{\perp T}GU^{\perp}<0 \tag{3.3.12}$$

同理有：

$$V^{\perp T}GV^{\perp}<0 \tag{3.3.13}$$

(2)充分性：

令R_{uv}是的$KerU\cap KerV$基，引入变量R_u和R_v使得

$$U^{\perp}=\begin{bmatrix}R_{uv} & R_u\end{bmatrix}$$

和

$$V^{\perp}=\begin{bmatrix}R_{uv} & R_v\end{bmatrix}$$

分别是$KerU$和$KerV$的基。

于是$KerU\oplus KerV$的基为

$$\begin{bmatrix}R_{uv} & R_u & R_v\end{bmatrix}$$

由此可以形成R^n的一个基，

$$T=\begin{bmatrix}R_{uv} & R_u & R_v & W\end{bmatrix}$$

这样的矩阵$T=\begin{bmatrix}R_{uv} & R_u & R_v & W\end{bmatrix}$是非奇异的。(3.3.11)式左乘$T$右乘$T^T$有：

$$TGT^T+(TU)X(TV)^T+TVX^T(TU)^T<0 \tag{3.3.14}$$

将TU,TV依T分块有

$$TU=\begin{bmatrix}0 & 0 & U_1 & U_2\end{bmatrix} \tag{3.3.15}$$

$$TV=\begin{bmatrix}0 & V_1 & 0 & V_2\end{bmatrix} \tag{3.3.16}$$

其中，$\begin{bmatrix}U_1 & U_2\end{bmatrix},\begin{bmatrix}V_1 & V_2\end{bmatrix}$为行满秩。

$$\begin{bmatrix}U_1^T\\ U_2^T\end{bmatrix}X\begin{bmatrix}V_1 & V_2\end{bmatrix}=\begin{bmatrix}X_{11} & X_{12}\\ X_{21} & X_{22}\end{bmatrix} \tag{3.3.17}$$

$$TGT^T=\begin{bmatrix}G_{11} & G_{12} & G_{13} & G_{14}\\ G_{12}^T & G_{22} & G_{23} & G_{24}\\ G_{13}^T & G_{23}^T & G_{33} & G_{34}\\ G_{14}^T & G_{24}^T & G_{34}^T & G_{44}\end{bmatrix} \tag{3.3.18}$$

于是，(3.3.14)式变为

$$\begin{bmatrix}G_{11} & G_{12} & G_{13} & G_{14}\\ G_{12}^T & G_{22} & G_{23}+X_{11}^T & G_{24}+X_{21}^T\\ G_{13}^T & G_{23}^T+X_{11} & G_{33} & G_{34}+X_{12}\\ G_{14}^T & G_{24}^T+X_{21} & G_{34}^T+X_{12}^T & G_{44}+X_{22}+X_{22}^T\end{bmatrix}<0 \tag{3.3.19}$$

其中，X_{ij}是任意的。由于$\begin{bmatrix}U_1 & U_2\end{bmatrix},\begin{bmatrix}V_1 & V_2\end{bmatrix}$为行满秩，因而有

$$X=\begin{bmatrix}U_1^T\\ U_2^T\end{bmatrix}^+\begin{bmatrix}X_{11} & X_{12}\\ X_{21} & X_{22}\end{bmatrix}\begin{bmatrix}V_1 & V_2\end{bmatrix}^+ \tag{3.3.20}$$

由 Schur 补引理，(3.3.19)式等价于

$$\Pi=\begin{bmatrix}G_{11} & G_{12} & G_{13}\\ G_{12}^T & G_{22} & G_{23}+X_{11}^T\\ G_{13}^T & G_{23}^T+X_{11} & G_{33}\end{bmatrix}<0 \tag{3.3.21}$$

$$G_{44}+X_{22}+X_{22}^T-\begin{bmatrix}G_{14}\\ G_{24}+X_{21}^T\\ G_{34}+X_{12}\end{bmatrix}\Pi^{-1}\begin{bmatrix}G_{14}\\ G_{24}+X_{21}^T\\ G_{34}+X_{12}\end{bmatrix}<0 \tag{3.3.22}$$

(3.3.11)式可解等价于(3.3.22)式可解。然而(3.3.22)式又等价于

$$\begin{bmatrix}I & 0 & 0\\ -G_{12}^TG_{11}^{-1} & I & 0\\ -G_{13}^TG_{11}^{-1} & 0 & I\end{bmatrix}\Pi\begin{bmatrix}I & -G_{11}^{-1}G_{12} & -G_{11}^{-1}G_{13}\\ 0 & I & 0\\ 0 & 0 & I\end{bmatrix}<0 \tag{3.3.23}$$

亦即

$$\begin{bmatrix} G_{11} & 0 & 0 \\ 0 & G_{22}-G_{12}^TG_{11}^{-1}G_{12} & X_{11}^T+\Lambda^T \\ 0 & X_{11}+\Lambda & G_{33}-G_{13}^TG_{11}^{-1}G_{13} \end{bmatrix}<0 \tag{3.3.24}$$

其中，

$$\Lambda = G_{23}^T - G_{13}^T G_{11}^{-1} G_{12} \tag{3.3.25}$$

由于 X_{11} 的任意性，(3.3.24)式的可解性等价于

$$\begin{cases} G_{11}<0 \\ G_{22}-G_{12}^TG_{11}^{-1}G_{12}<0 \\ G_{33}-G_{13}^TG_{11}^{-1}G_{13}<0 \end{cases} \tag{3.3.26}$$

或者等价于

$$\begin{bmatrix} G_{11} & G_{12} \\ G_{12}^T & G_{22} \end{bmatrix}<0 \tag{3.3.27}$$

或者

$$\begin{bmatrix} G_{11} & G_{13} \\ G_{13}^T & G_{33} \end{bmatrix}<0 \tag{3.3.28}$$

代入 $U^{\perp}, V^{\perp}$ 及 G_{ij} 的定义，即可以得到(3.3.12)和(3.3.13)式。

[证毕]

3.4 控制理论中的 *LMI* 问题

3.4.1 正实引理

定义 3.4.1（正实函数定义） 对于有理函数 $F(s)$，当满足下列条件时，称其为正实函数。

(1) $F(s)$ 的极点只存在于包含虚轴的左半平面内，且虚轴上的极点为一阶极点，其留数为正值；

(2) 对于所有的 w 有，$\mathrm{Re}(F(jw)) \ge 0$。

正实函数是网络理论中无源回路的策动函数。

$\Sigma(A,b,c)$ 为正实系统的充分必要条件：设 $T(s)$ 为 $\Sigma(A,b,c)$ 的系统传递函数，$T(s)$ 为正实的充分必要条件为存在对称正定矩阵 P 以及相应维数的向量 l，满足下列二式

$$\begin{cases} A^TP+PA=-l^Tl \\ Pb=c^T \end{cases}$$

由上述的正实函数的定义可以得到单输入单输出系统的正实引理。

引理 3.4.1 给定系统实现 $\Sigma(A,b,c)$，其传递函数为 $T(s)=c(sI-A)^{-1}b$。当 $T(s)$ 为正实的，即

$$\mathrm{Re}\,T(s)>0\text{，对于所有的}\,\mathrm{Re}(s)>0 \tag{3.4.1}$$

当且仅当存在对称实矩阵 $P=P^T>0$ 满足

$$A^T P + PA < 0\text{，且 } Pb = c^T \tag{3.4.2}$$

定义 3.4.2（正实矩阵定义）设$F(s)$为$r \times r$的实有理传递函数矩阵，如果$F(s)$满足下列两个条件，则称$F(s)$为正实矩阵。

（1）$F(s)$的各元素的极点只存在于包含虚轴的左半平面内，但虚轴上的极点是一阶极点，其留数矩阵为正半定矩阵；

（2）对于所有的w，下式成立

$$F(jw) + F^T(-jw) \ge 0$$

有理正实矩阵系统理论的充分必要条件：设$F(s)$为$r \times r$的实有理传递函数矩阵，同时设它的一个最小实现为$\Sigma(A,B,C,D)$。使$F(s)$为正实矩阵的充分必要条件为存在对称正定矩阵P和适当维数的矩阵W_0和L，同时满足以下三式：

$$\begin{cases} A^T P + PA = -L^T L \\ PB = C^T - L^T W_0 \\ {W_0}^T W_0 = D + D^T \end{cases}$$

由上述的正实矩阵的定义可以得到多输入多输出系统的正实引理。

引理 3.4.2　给定系统实现$\Sigma(A,B,C,D)$，A的特征值都具有负实部，(A,B)可控，并且$D + D^T > 0$。下列陈述等价：

①系统$\Sigma(A,B,C,D)$其状态空间表述

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \\ x(0) = 0 \end{cases} \tag{3.4.3}$$

该系统是 Passive 的，即对于所有的控制u以及时间$T \ge 0$，满足

$$\int_0^T u(t)^T y(t)dt \ge 0 \tag{3.4.4}$$

②　系统$\Sigma(A,B,C,D)$的传递函数矩阵$T(s) = C(sI - A)^{-1}B + D$，是正实的，即对于所有的$\mathrm{Re}(s) > 0$，均有

$$T(s) + T^H(s) > 0 \tag{3.4.5}$$

③对于LMI

$$\begin{bmatrix} A^T P + PA & PB - C^T \\ B^T P - C & -(D + D^T) \end{bmatrix} < 0 \tag{3.4.6}$$

存在对称矩阵$P = P^T > 0$使之可解。

④存在矩阵$P = P^T$，满足

$$A^T P + PA + (PB - C^T)(D + D^T)^{-1}(PB - C^T)^T = 0 \tag{3.4.7}$$

3.4.2 有界实引理

引理 3.4.3　给定系统实现 $\Sigma(A,b,c)$，其传递函数为 $T(s)=c(sI-A)^{-1}b$。当 $T(s)$ 为有界实的，即

$$\|T(s)\|_\infty<\gamma\text{，}A\text{稳定，即}\operatorname{Re}(\lambda_i(A))<0 \tag{3.4.8}$$

当且仅当存在对称实矩阵 $P=P^T>0$ 满足

$$A^TP+PA<0\text{，且}Pb=c^T \tag{3.4.9}$$

引理 3.4.4　给定系统实现 $\Sigma(A,B,C,D)$，A 的特征值都具有负实部，(A,B) 可控，并且 $D+D^T>0$。下列陈述等价：

① 系统 $\Sigma(A,B,C,D)$ 其状态空间表述

$$\begin{cases}\dot{x}=Ax+Bu\\ y=Cx+Du\\ x(0)=0\end{cases} \tag{3.4.10}$$

该系统是 Passive 的，即对于所有的控制 u 以及时间 $T\geq 0$，满足

$$\int_0^T u(t)^T y(t)dt\geq 0 \tag{3.4.11}$$

② 系统 $\Sigma(A,B,C,D)$ 的传递函数矩阵 $T(s)=C(sI-A)^{-1}B+D$，是正实的,，即对于所有的 $\operatorname{Re}(s)>0$，均有

$$\|T(s)\|_\infty<\gamma\text{，}A\text{稳定，即}\operatorname{Re}(\lambda_i(A))<0 \tag{3.4.12}$$

③ 存在对称矩阵 $P=P^T>0$ 满足 LMI

$$\begin{bmatrix}A^TP+PA & PB & C^T\\ B^TP & -\gamma I & D^T\\ C & D & -\gamma I\end{bmatrix}<0 \tag{3.4.13}$$

④ 存在矩阵 $P=P^T$，满足

$$A^TP+PA+(PB-C^T)(D+D^T)^{-1}(PB-C^T)^T=0 \tag{3.4.14}$$

对于离散系统有类似的定理。

3.4.3　*LMI* 与Riccati方程、Riccati不等式以及H∞范数的关系

设系统 $\Sigma(A,B,C,D)$ 的传递函数

$$T(s)=C(sI-A)^{-1}B+D \tag{3.4.15}$$

且系统矩阵 A 为稳定阵以及 $\overline{\sigma}(D)<1$。则

$$T\|P(s)\|_\infty<1 \tag{3.4.16}$$

对于下述的 *LMI*：

$$\begin{bmatrix}A^TX+XA+C^TC & XB+C^TD\\ B^TX+D^TC & D^TD-I\end{bmatrix}<0 \tag{3.4.17}$$

可解的充分必要条件是如下 Riccati 方程

$$X(A+BR^{-1}D^TC)+(A+BR^{-1}D^TC)^TX+XBR^{-1}B^TX$$
$$+C^T(I+DR^{-1}D^T)C=0 \qquad (3.4.18)$$

具有半正定解 $X\geq 0$，且 $A+BR^{-1}(B^TX+D^TC)$ 是稳定矩阵，其中，$R=I-D^TD>0$。

这个定理是有界实引理的直接结果，也是最著名的H∞范数定理之一。下面给出的Riccati不等式定理与上述定理等效。

设系统的传递函数 $T(s)=C(sI-A)^{-1}B+D$，且系统矩阵 A 为稳定阵以及 $\bar{\sigma}(D)<1$。则

$$\|T(s)\|_\infty<1 \qquad (3.4.19)$$

式(3.4.17)可解的充分必要条件是如下 Riccati 不等式具有正定解 $X>0$。

$$X(A+BR^{-1}D^TC)+(A+BR^{-1}D^TC)^TX+XBR^{-1}B^TX$$
$$+C^T(I+DR^{-1}D^T)C<0 \qquad (3.4.20)$$

Kalman 和 Bucy 首先研究了连续时不变系统 Riccati 方程解的收敛性问题，并得出在 (A,B) 可控和 (A,C) 可测条件下，Riccati 方程解以指数速度收敛到唯一的对称正定解。后来这一条件又被放宽为 (A,C) 可检测和 (A,B) 可稳定。这些结果一度被认为是 Riccati 方程有稳定解的充要条件。

上述的 Riccati 方程和 Riccati 不等式对我们以后的工作具有很大帮助，并且它们是研究线性矩阵不等式的基础。

3.5 *MATLAB LMI-tools* 求解 *LMI* 可行性问题方法

MatLab 中的线性矩阵不等式(*LMI*)工具箱是求解一般 *LMI* 问题的一个高性能软件包。由于它面向结构的 *LMI* 表示方式，使得各种 *LMI* 都能够以自然块矩阵地形式加以描述。一个 *LMI* 问题一旦确定，就可以通过调用适当的 *LMI* 求解器进行数值求解。

LMI 工具箱提供了确定、处理和数值求解 *LMI* 的工具，主要功能有：*LMI* 的描述、获取 *LMI* 得信息、修改 *LMI*、三个一般 *LMI* 的求解器和验证计算结果。

3.5.1 线性矩阵不等式的描述和确定

LMI 工具箱可以处理的 *LMI* 问题的格式为：

$$N^TL(X_1,\cdots,X_N)N<M^TR(X_1,\cdots,X_N)M$$

其中 N 和 M 是外因子，具有相同维数的给定矩阵。$X_1,\cdots,X_N$ 是具有一定结构的变量，它们可以是矩阵、向量和标量。必须注意的是，在一个 *LMI* 描述中，左边总是指不等式较小的一边。

一个 *LMI* 不论多么复杂，都可以通过描述其中每一块的各项内容来确定这个 *LMI*，一般情况需要一下两个步骤。

(1) 给出每个矩阵变量 $X_1,\cdots,X_N$ 的结构和维数；

(2) 描述每个 *LMI* 中各个项的内容。

这个过程产生所描述 *LMI* 系统的一个内部表示，它以一个单一向量的形式储存在计算机内，通常用一个名字命名，例如 lmisys。该内部表示可以在处理这个 *LMI* 的程序中调用。

下面先介绍几个最常使用的函数的功能和格式。

(1) Setkmis 和 getkmis

一个 *LMI* 系统描述以 setlmis 开始，以 getkmis 结束。当要确定一个新的 *LMI* 系统时，输入

```
setlmis([ ])
```

如果需要将一个 *LMI* 添加到一个名为 lmiso 的现有 *LMI* 系统时，则输入

```
setlmis(lmis0)
```

当一个 *LMI* 系统被完全确定好后，输入

```
lmisys=getlmis
```

该命令返回这个 *LMI* 的内部表示 lmisys。

(2) lmivar

函数 lmivar 用来描述出现在 *LMI* 系统中的矩阵变量，每一次只能描述一个矩阵变量。矩阵变量的描述包括该矩阵变量的结构。该函数的输入格式为：

```
X=lmivar(type,struct)
```

它定义了一个新的矩阵变量 X 。第一个输入量 type 确定了矩阵变量 X 的类型，类型一般分为三类：type＝1，对称块对角结构；type＝2，长方形结构；type＝3，其它结构。

Struct 是一个和变量 X 有相同维数的矩阵，其中的每一个元取值为：

$$\text{struct}(i,j)=\begin{cases} 0, & \text{如果}X(i,j)=0 \\ n, & \text{如果}X(i,j)=x_n \\ -n, & \text{如果}X(i,j)=-x_n \end{cases}$$

(3) lmiterm

在确定了矩阵变量之后，还需要确定每一个 *LMI* 中的各项的内容。*LMI* 的项指构成这个 *LMI* 地块矩阵中的求和项。一般分成三类：常数项，变量项合外因子。

用命令 lmiterm 每次可以确定 *LMI* 的一个项的内容。例如，对 *LMI*

$$\begin{bmatrix} A^TX+XA+C^TSC & XB \\ B^TX & S \end{bmatrix}$$

可以用以下一组命令来描述：

```
lmiterm([1 1 1 X],1,A,'s')
lmiterm([1 1 1 S],C',C)
lmiterm([1 1 2 X],1,B)
lmiterm([1 2 2 s],-1,1)
```

在每一条命令中，第一项是一个 4 元向量，它说明了所描述的项所在地位置和特征，其中：

¶ 第一个元表示所描述的项属于那一个 *LMI* 。值 m 表示第 m 个不等式的左边，值－m 表示第 m 个不等式的右边。该项也可以直接用不等式的名，例如 lmi1=newlmi 中的 lmi1。

¶ 第 2 和第 3 个元表示所描述的项所在块的位置。例如[1 1 2 1]表示所描述的项位于第一个不等式左边内因子的块(1,2)中。第 2 和第 3 个元素均取零表示所描述的项在外因子中。

¶ 最后一个元素说明所描述的项是常数项还是变量项。0 表示常数项。

Lmiterm 第 2 项合第 3 项包含了变量和数据的左右系数。第 4 项是可选项，且只能是's'。

3.5.2 线性矩阵不等式的求解器

LMI 工具箱提供了用于求解以下三个问题的 *LMI* 求解器。

(1) 可行性问题: [tmin,xfeas]=feasp(lmisys,options,target)

它是通过求解如下的一个辅助凸优化问题

$$\begin{aligned}&\min t\\ &s.t \quad A(x)-B(x)\le tI\end{aligned}$$

来求解 *LMI* 系统 lmisys 的可行性问题。其全局最优值用 tmin 表示，如果 tmin<0，则系统 lmisys 是可行的，此时 feasp 输出的第二个参量 xfeas 给出了该 *LMI* 系统决策变量的一个可行解。进而用命令 dec2mat 可以得到系统 lmisys 矩阵变量的具体解。

(2) 线性规划问题：[copt,xopt]=mincx(lmisys,c,options,xinit,target)

具有 *LMI* 约束的一个线性目标函数的极小化问题

$$\begin{aligned}&\min_x c^T x\\ &s.t \quad A(x)<B(x)\end{aligned}$$

相应的求解器是 mincx。同样问题中的 *LMI* 系统由 lmisys 表示。向量 c 和决策变量 x 有相同的维数。对于由矩阵变量表示的线性目标函数，可以应用函数 defcx 来得到适当地向量 c。函数 mincx 返回到目标函数 $c^T x$ 的全局最优值 copt 和决策变量的最优解 xopt，相应的矩阵变量的最优值可以应用函数 dec2mat 从 xopt 得到。

函数 mincx 的输入量中除了 lmisys 和 c 以外，其它的输入是可选择的，具体应用可以参考 *LMI* 工具箱的使用手册。

(4) 广义特征值的极小化问题：[lopt,xopt]=gevp(lmisys,nlfc,options,linit,xinit,target)

求解器 gevp 是通过求解如下的广义特征值的极小化问题

$$\begin{aligned}&\min_x \lambda\\ s.t.\quad &C(x)<D(x)\\ &0<B(x)\\ &A(x)<\lambda B(x)\end{aligned}$$

如果问题的 *LMI* 约束是可行的，则 gevp 给出了优化问题的全局最小值 lopt 和决策向量 x 的最优解 xopt。相应的矩阵变量的最优解可以应用 dec2mat 得到。

输入分量 lmisys 表示当 $\lambda=1$ 时的 *LMI* 系统。包含 λ 的 *LMI* 系统称为线性分式约束。线性分式约束的个数用 nlfc 表示。其它输入分量都是可选择的。

在调用求解器 gevp 时，须遵循以下原则：

¶ 确定包含 λ 的线性矩阵不等式：$A(x)<B(x)$（注意没有 λ 的不等式）;

¶ 总是把 $A(x)<B(x)$ 放在线性矩阵不等式系统的最后;

¶ 要求有约束 $0<B(x)$，或者保证 $0<B(x)$ 成立的任何其它约束。